

# Model for the hydration of non-polar compounds and polymers

Pierpaolo Bruscolini\*

*Istituto Nazionale per la Fisica della Materia (INFM) and Dipartimento di Fisica, Politecnico di Torino,  
Corso Duca degli Abruzzi 24, I-10129 Torino, Italy*

Lapo Casetti†

*Istituto Nazionale per la Fisica della Materia (INFM), Unità di Ricerca di Firenze,  
Largo Enrico Fermi 2, I-50125 Firenze, Italy  
(February 1, 2008)*

We introduce an exactly solvable statistical-mechanical model of the hydration of non-polar compounds, based on grouping water molecules in clusters where hydrogen bonds and isotropic interactions occur; interactions between clusters are neglected. Analytical results show that an effective strengthening of hydrogen bonds in the presence of the solute, together with a geometric reorganization of water molecules, are enough to yield hydrophobic behavior. We extend our model to describe a non-polar homopolymer in aqueous solution, obtaining a clear evidence of both “cold” and “warm” swelling transitions. This suggests that our model could be relevant to describe some features of protein folding.

PACS number(s): 05.20.-y; 05.40.Fb; 61.25.Hq; 87.10.+e

## I. INTRODUCTION

The physical properties of liquid water, despite the extensive studies devoted to their investigation, are not yet completely understood. The behavior of water when nonpolar solutes are present is still a matter of debate, even if simplified models [1,2] as well as numerical simulations (see Ref. [3] and references quoted therein) have shed some light on it. The hydration of non-polar compounds has a free energy cost, resulting in a net “attraction” among solute molecules which tend to aggregate in order to minimize exposition to water (“hydrophobic effect”). More precisely, upon hydration of a single molecule from the gaseous phase, the free energy and specific heat changes  $\Delta F$  and  $\Delta C$  are positive over the entire temperature range accessible to experiments; the internal energy and entropy differences  $\Delta E$  and  $\Delta S$  are negative for low temperatures, then become positive when the temperature is raised. The two temperatures at which  $\Delta E$  and  $\Delta S$  are zero are commonly referred to as  $T_H$  and  $T_S$ , respectively. The origin of the lowering of the entropy for  $T < T_S$  has been related to the ordering of water molecules around the solute to preserve bonds [4,5] as well as to the opening of a “cavity” in water to settle the solute, without ordering of water molecules [6]; but the question is still controversial. Hydrophobicity plays a fundamental role in many physical processes, among which the long-standing problem of protein folding, since the hydrophobic effect is believed to be one of the most relevant “forces” guiding the protein into the folded native state, i.e., the unique spatial structure in which the protein is biologically active [7]. At present, the whole folding process can be studied only using coarse-grained models (all-atom simulations of proteins and solvent molecules are feasible only on timescales several order of magnitude smaller

than the folding time). In many models water is not explicitly considered and solvent-induced aggregation is described through temperature-independent effective attraction among nonpolar monomers. An important feature of real (globular) proteins is that the native state becomes unstable not only at high temperature (*warm unfolding*), but also at low temperatures (*cold unfolding*) [8]; cold swelling has also been detected in some homopolymers [9,10]. A major drawback of the above simplified models is that they cannot account for cold unfolding, because there the native (folded) state is identified with the model ground state, so that the protein is folded at low temperatures down to  $T = 0$ .

In this paper we propose a “minimal” model capable of reproducing the thermodynamic hallmarks of hydrophobicity in single molecule’s hydration, within the framework of the “water ordering” hypothesis; moreover, we keep it as simple as to be analytically integrable, in order to apply it to polymer hydration. For its simplicity, the model is not intended to reproduce all the details of hydrophobic behavior: rather it gives a correct overall description of the phenomenology and allows to understand what are the key ingredients that produce it.

In the case of a nonpolar homopolymer we recover both “cold” and “warm” swelling transitions, thus strengthening the idea, already put forward in Refs. [12], that an explicit, though simplified, description of water can also provide a framework for a unified treatment of both transitions in proteins.

## II. THE MODEL

We describe water molecules as two-dimensional objects with three hydrogen-bonding arms, namely, the planar projection of the tetrahedral coordination of water

molecules, as in Refs. [13,5]. Simulations show that such a crude model reproduces the main features of water [5]. The molecules in contact with a solute will be referred to as the *hydration shell*. Hydrophobicity is a collective phenomenon, yet experiments [8] show additivity in the nonpolar area exposed to the solvent. Hence, there exists a cluster scale above which correlations may be neglected. We consider the hydration shell to be made up of  $z$  such clusters, each of  $m$  interacting molecules, and we disregard interactions between clusters. We assume that the ground state of a cluster is characterized by a completely formed hydrogen-bond network, both in the bulk and in the hydration shell case. Thus each molecule in the ground state has on average  $3/2$  hydrogen bonds. Hydrogen-bond energies are very sensitive to the geometry of the bond, and the latter will be affected by the presence of a solute, so that we assume that bond energies may differ in the “bulk” and “shell” cases. We model this by simply assuming that there is an energy difference (per molecule)  $K$  between bulk and shell ground states, given by

$$K = \frac{3}{2} (h_b - h_s) + J, \quad (1)$$

where  $h_b$  and  $h_s$  are the hydrogen-bond energies for bulk and shell molecules, respectively. The fact that  $h_b \neq h_s$  is not, in principle, the unique possible cause of an energy shift in the ground states: not only a solute perturbs the water hydrogen bonds, but also has a different direct interaction with water. To take into account this kind of contributions to the ground-state energy difference, assuming that they can be reasonably averaged on the molecular scale, we introduce the constant  $J$  in Eq. (1) above.

### A. Partition function of the hydration shell

Our goal is to model the thermodynamics of hydration, hence we need to evaluate the partition function for the cluster in both cases:

$$\mathcal{Z}_{\bullet}^{\text{clu}} = \int_0^\infty d\varepsilon g_{\bullet}(\varepsilon) e^{-\beta \mathcal{H}_{\bullet}^{\text{clu}}(\varepsilon)}, \quad (2)$$

where  $\bullet = b, s$  (bulk or shell, respectively). Here,

$$\mathcal{H}_{\bullet}^{\text{clu}} = \varepsilon + Km \delta_{\bullet,s}, \quad (3)$$

and  $g_{\bullet}(\varepsilon)$  is the density of states of the cluster at the energy  $\varepsilon$  above its ground state: it encodes all the important features determining the system thermodynamics. How can we estimate  $g_{\bullet}(\varepsilon)$ ? Given an energy  $\varepsilon$  to the cluster we expect to have a certain number of broken bonds and assume equipartition of the energy on the various degrees of freedom of the cluster. Moreover, bond breaking allows molecules to access a wider configuration space: while the ground state is unique, different

geometrical arrangements of the molecules can have the same number of bonds. To describe it simply enough, we take each bond to be represented by a harmonic potential, and to be broken independently of the others. Thus, at the cluster energy  $\varepsilon$  there will be a probability

$$p_{\bullet}(\lambda, \varepsilon) = \binom{\frac{3}{2}m}{\lambda} p_{1\bullet}(\varepsilon)^\lambda (1 - p_{1\bullet}(\varepsilon))^{\frac{3}{2}m - \lambda} \quad (4)$$

to break  $\lambda$  bonds. Here  $p_{1\bullet}(\varepsilon)$  is the probability that a bond acquires an energy larger than  $h_{\bullet}$  and breaks, when  $\varepsilon$  is equipartitioned on  $D$  degrees of freedom:  $p_{1\bullet}(\varepsilon) = e^{-h_{\bullet}/\tilde{T}}$  and  $\tilde{T} = 2\varepsilon/D$ . For the sake of simplicity we assume  $D = 2mf$  ( $f$  are the degrees of freedom of one molecule:  $f = 3$  in two dimensions) as if molecules were at least singly bonded at all energies. In the same spirit we take the density of states  $\omega(\lambda, \varepsilon)$  of the system of harmonic oscillators, resulting when  $\lambda$  bonds are broken, to be independent of  $\lambda$ :  $\omega(\lambda, \varepsilon) \simeq \omega(\varepsilon) = \mathcal{C} \varepsilon^{\frac{D}{2}-1}$ , where  $\mathcal{C}$  is a constant. To account for the geometric degeneracy we assume just two orientations for each molecule, (E and I; see Fig. 1), and give an estimate of  $\gamma_{\bullet}(\nu)$ , the number of accessible geometric arrangements of the cluster when  $\lambda = 3m/2 - \nu$  bonds are broken. Then, our ansatz for  $g_{\bullet}(\varepsilon)$  will be the sum, over  $\lambda$ , of the probability  $p_{\bullet}(\lambda, \varepsilon)$  of breaking  $\lambda$  bonds, times the density of states  $\omega(\varepsilon)$  of each geometric configuration, times the number  $\gamma_{\bullet}(\nu)$  of configurations at  $\lambda$  broken bonds:

$$g_{\bullet}(\varepsilon) = \sum_{\lambda=0}^{\frac{3}{2}m} p_{\bullet}(\lambda, \varepsilon) \omega(\varepsilon) \gamma_{\bullet} \left( \frac{3m}{2} - \lambda \right). \quad (5)$$

In the estimation of  $\gamma_{\bullet}(\nu)$  the bookkeeping is easier if we attribute each bond to just one of the two bonding molecules. Each cluster can have no more than  $3m/2$  bonds: to ensure this, we take  $m$  even and fictitiously split the cluster in two equal sets,  $a$  and  $b$ . Group  $a$ -molecules will form at most one bond: a shell one when in state-E and an external one when in state-I. Group  $b$ -molecules form at most two bonds: the external and a shell one in state-E, an external and the internal one in state-I. Internal bonds will be possible only for ( $b$ ,I) molecules, if they find a bonding partner. The number of internal bonding sites,  $q_{\bullet}$ , will distinguish “shell” from “bulk” cases: in the former  $q_s = 0$ , while in the latter  $0 < q_b \leq m$ . In fact, due to the geometry of the internal water molecules, possibly not all the  $m$  positions allow internal bonds, even in the bulk case. State-E molecules can form one shell bond, while state-I cannot: hence, the probability of a configuration with  $s$  state-I molecules separating  $(m-s)$  state-E ones, so that  $k$  shell bonds are forbidden and  $\chi = m - s - k$  can be formed, is [14]

$$\pi(s, k) = 2^{-m} \binom{s}{k} \binom{m-s}{k}. \quad (6)$$

The probability that one of these configurations also has  $i$  internal bonds depends, first, on the probability of fishing out  $j$  ( $b$ ,I)-molecules among the  $s$  in state-I, given

the total number of molecules  $m$  and of  $b$ -molecules  $m/2$ :  $\pi_h(j; s, m/2, m)$ , where  $\pi_h$  is a hypergeometric probability, i.e.,

$$\pi_h(s; n, S, N) = \binom{S}{s} \binom{N-S}{n-s} \left[ \binom{N}{n} \right]^{-1}. \quad (7)$$

Then, one must consider the probability of placing  $i$  of the  $j$  molecules with internal arms, in the  $q_\bullet$  good positions for bond formations, on a total of  $m$  possibilities:  $\pi_h(i; j, q_\bullet, m)$ . The product of all the above probabilities gives the fraction of conformations that has  $s$  state-I molecules and is able to form  $\chi$  shell bonds,  $i$  internal ones and  $m/2 + s - j$  external ones (assuming that all the external arms form bonds). If we now let the bonds be also broken, we must notice that any pattern geometrically allowing more than  $\nu$  bonds will also contribute to  $\gamma_\bullet(\nu)$ : we just need to consider the excess bonds as broken – we choose them among the external, internal and shell ones with a binomial probability  $\pi_b(n, N) \equiv \pi_b(n, N, 1/2)$ , where

$$\pi_b(n, N, p) = \binom{N}{n} p^n (1-p)^{(N-n)}. \quad (8)$$

Hence, upon summing over all the geometric arrangements that can contribute to a pattern with  $\nu$  bonds, we finally get

$$\gamma_\bullet(\nu) = \sum_{\nu_s=0}^{\nu} \sum_{\nu_e=0}^{\nu-\nu_s} \sum_{s=0}^m \sum_{k=0}^{k_{\max}} \sum_{j=0}^s \sum_{i=0}^j \xi_\bullet(s, k, j, i, \nu_e, \nu_s, \nu_i), \quad (9)$$

where  $\nu_i = \nu - \nu_s - \nu_e$ ,  $k_{\max} = \min(s, m - s - \nu_s)$  and

$$\begin{aligned} \xi_\bullet(s, k, j, i, \nu_e, \nu_s, \nu_i) &= \pi(s, k) \pi_h\left(j; s, \frac{m}{2}, m\right) \pi_h(i; j, q_\bullet, m) \\ &\times \pi_b(\nu_i, i) \pi_b\left(\nu_e, \frac{m}{2} + s - j\right) \pi_b(\nu_s, m - s - k). \end{aligned} \quad (10)$$

The partition function can now be evaluated from Eq. (2) and Eqs. (5,4,9). We get:

$$\mathcal{Z}_\bullet^{\text{clu}} = A_\bullet e^{-\beta K m \delta_{\bullet, s}} \quad (11)$$

where  $A_\bullet = \mathcal{C} \sum_{\lambda=0}^{\frac{3}{2}m} \left(\frac{\frac{3}{2}m}{\lambda}\right) \gamma_\bullet(3m/2 - \lambda) I_\bullet(\lambda)$ ,  $\mathcal{C}$  is a constant and

$$I_\bullet(\lambda) = \frac{\left(\frac{D}{2}-1\right)!}{\beta^{\frac{D}{2}}} \delta_{\lambda,0} + 2 \sum_{j=\delta_{\lambda,0}}^{\lambda'} (-1)^j \binom{\lambda'}{j} \sigma_\bullet^{\frac{D}{2}} K_{\frac{D}{2}}(\tau_\bullet), \quad (12)$$

with  $K_D(\tau_\bullet)$  the Bessel- $K$  function,  $\lambda' = 3m/2 - \lambda$ , and

$$\sigma_\bullet = (2\beta)^{-1} D h_\bullet(j + \lambda), \quad \tau_\bullet = \sqrt{2\beta h_\bullet D(j + \lambda)}. \quad (13)$$

## B. Statistical mechanics of non-polar hydration

We now consider the transfer of a nonpolar molecule from its gas phase to water (i.e., hydrophobic hydration). We describe it by taking  $z$  clusters, of  $m$  water molecules, in the bulk case and substituting the internal water with the solute. The free energy change  $\Delta F$  is given by

$$\beta \Delta F = -z \log x, \quad (14)$$

with  $x = A_s A_b^{-1} \exp(-\beta K m)$ . Similar expressions hold for energy, entropy and specific heat changes. The temperature dependence of these functions, reported in Fig. 2, shows the hallmarks of hydrophobic behavior: we find a maximum in the free energy cost and minima in both  $\Delta E$  and  $T\Delta S$ ; as  $T$  grows, first  $\Delta E$  and then  $T\Delta S$  cross the zero – at  $T_H$  and  $T_S$ , respectively – and eventually become positive. Moreover, we correctly find a pronounced and positive peak in the specific heat difference  $\Delta C$ . This cannot be found within the information theory approximation to the cavity model: see Ref. [15]. It turns out that both a ratio  $h_s/h_b > 1$  and the difference between  $\gamma_b(\nu)$  and  $\gamma_s(\nu)$  concur to produce the correct thermodynamic behavior: with  $h_s = h_b$  the difference between  $\gamma_b(\nu)$  and  $\gamma_s(\nu)$  is not sufficient to produce the minimum in  $\Delta E$ , while  $\gamma_b(\nu) = \gamma_s(\nu)$  (i.e., same chance of internal bonds,  $q_b = q_s$ ) yields too small a peak in  $\Delta C$ .

Then, according to our model, hydrophobicity appears as related to the average strengthening of hydrogen bonds, together with geometric pattern changes, upon the introduction of a nonpolar solute.

In particular, having tighter shell bonds than bulk ones ( $h_s > h_b$ ) appears to be essential, in our model, to get a negative minimum in  $\Delta E$ . Indeed, even starting from equal energy ground states in the bulk and shell cases (i.e.,  $K = 0$ ), if  $h_s > h_b$  it will be easier to break bonds in bulk than in shell, according to Eq. (4). This yields that at “intermediate” temperatures shell water will have lower energy than bulk, i.e.  $\Delta E < 0$ .

The predictions on the trends of thermodynamic functions are in qualitative agreement with experimental results for noble gases (see the results by Crovetto and coworkers [16] as reported in [5]) and for nonpolar surfaces in proteins [8]. They are also in agreement with the Monte Carlo results for the two-dimensional “MB” water model in Ref. [5], where an average strengthening of shell hydrogen bonds is detected, too, as well as negative values of  $\Delta C$  at low temperatures.

It should be mentioned, though, that negative  $\Delta E$  and  $\Delta S$ , and positive  $\Delta F$  can be obtained for the MB model within the information theory approximation applied to the cavity opening in pure water Ref. [15], even if, as already stated, this approach cannot recover the correct trends for the specific heat. Since cavity opening disregards bond differences in bulk and shell, this fact would imply that hydrogen bond differences are irrelevant to the subject.

It is likely that both cavity formation and hydrogen-bond differences contribute to the hydrophobic hydration, and an ultimate answer will come from simulations in three dimensions when they will become accurate enough as to predict specific heats. As far as the present model is concerned, we consider the good overall agreement with experiments and two-dimensional simulations as satisfactory to justify the use of this model in polymer hydration studies.

### C. Cold and warm unfolding of a non-polar homopolymer

Let us now turn to the study of a nonpolar homopolymer in solution, taking into account just the behavior of water clusters in the vicinity of a monomer, and disregarding interactions between monomers and between water clusters. Our goal is, in fact, to understand the effect of the hydrophobic effect alone on polymer behavior. We model a polymer as a  $N$ -step self-avoiding walk (SAW) on a two dimensional lattice with coordination  $z$ . On each lattice site there can be either a monomer or  $z$  clusters of  $m$  water molecules, so that each monomer-water contact involves one cluster. The Hamiltonian follows from Eq. (3):

$$\mathcal{H} = \sum_{j=1}^{N_W} \sum_{\mu=1}^z \varepsilon_{j\mu} + K m l_j, \quad (15)$$

where  $N_W = (z-2)N + 2$  is the highest number of water sites that can be in contact with the polymer, and  $l_j$  is the number of contacts between the  $j$ -th water site and the monomers. The partition function of the polymer in solution reads as:

$$\mathcal{Z} = \sum_C Z(C) = \sum_{n_c} \zeta(n_c) Z(n_c) \quad (16)$$

where  $C$  are the conformations of the polymer and  $Z(C)$  the restricted partition function, obtained tracing over water variables at fixed conformation  $C$ . Due to the form of the Hamiltonian given in Eq. (15),  $Z(C)$  depends only on the total number of water-monomer contacts  $n_c$ , and  $\zeta(n_c)$  is the number of SAWs characterized by the same value of  $n_c$ .  $\mathcal{Z}$  can be factorized as  $\mathcal{Z} = \mathcal{Z}_b \mathcal{Z}_I$ , where  $\mathcal{Z}_b = (\mathcal{Z}_b^{\text{clu}})^{zN_W} = A_b^{zN_W}$  is the contribution of all water sites when in contact to other water, and  $\mathcal{Z}_I = \sum_{n_c} \zeta(n_c) x^{n_c}$ . According to the above factorization, the specific heat is the sum of a bulk contribution  $C_b$  and of an interaction contribution  $C_I$ . We introduce also the average number of water-monomer contacts,  $\langle n_c \rangle = x \frac{\partial}{\partial x} \log \mathcal{Z}_I$ , which is a measure of the compactness of the polymer. To exactly evaluate the above quantities,  $\zeta(n_c)$  should be obtained through an exhaustive enumeration of the SAWs. However, if we restrict ourselves to a square lattice, the numerical estimates reported in [17] allow us to write

$$\zeta(n_c) \simeq \zeta_0 \frac{1}{w(n_c)!} (\alpha_0 N)^{w(n_c)} \exp(\alpha_0 N), \quad (17)$$

where  $w(n_c) = (N_W - n_c)/2$  is the number of monomer-monomer contacts,  $\alpha_0 = 0.164$  and  $\zeta_0$  is the number of SAWs of length  $N$ . Such an estimate is expected to be very good if  $N$  is large [17]. Hence, an analytical expression can be found for  $\mathcal{Z}_I$ :

$$\mathcal{Z}_I = \frac{1}{w_{\max}!} x^{N_W} \Gamma(w_{\max} + 1, \frac{\alpha_0 N}{x^2}) e^{-\alpha_0 N(1 - \frac{1}{x^2})}, \quad (18)$$

where  $\Gamma(n, x)$  is the incomplete  $\Gamma$ -function and  $w_{\max} = (N_W - n_c^{\min})/2$  (the smallest number of contacts is the perimeter of the globule,  $n_c^{\min} \simeq 2\sqrt{N\pi}$ ). The results for  $C_I$  and  $\langle n_c \rangle$  are reported in Fig. 3.

The presence of both “cold” and “warm” collapse transitions, signalled by the drop of the number of contacts and by the jumps in the specific heat is strikingly evident. Notice that the polymer swells at low temperature even with  $K = 0$ : water-clusters’ energy at  $T = 0$  is not affected by the contact with monomers, and cold unfolding is not a trivial effect of a swollen-ground-state dominance. As  $N$  grows both transitions get sharper, thus suggesting the existence of true phase transitions in the thermodynamic limit. The above phenomenology is very close to that of proteins in solution [8]. Moreover, swelling at low temperatures has been found in homopolymers like poly(N-isopropylacrylamide) (PNIPAM) and poly(N-isopropylmethacrylamide) (PNIPMAM) [9], while warm swelling has been detected for a number of homopolymers in non-aqueous solution [18].

Both PNIPAM and PNIPMAM present a polar peptide group in the side chain and hence are able to form hydrogen bonds; hence, one could object that it is not appropriate to propose a comparison with their phenomenology, since polar groups favourably interact with water and help the swelling of the chain. Actually, we are not aware of any calorimetric experiment on completely non-polar homopolymers in water (that would represent ideal tests to our predictions, provided that they swell in the range  $T = 0 - 100^\circ\text{C}$  and the solution is so dilute that aggregation is negligible). However, hydrogen bonds can be formed both in the coil phase, with water, and in the globule one (and are indeed considered a possible cause for the detected hysteresis in the coil-globule transition [10]). Assuming a negligible overall contribution of hydrogen bonds to the stability of the coil over the globule phase, one can admit that the cold transition in PNIPAM and PNIPMAM is mainly due to the temperature dependence of the free-energy of hydration of the nonpolar groups. This is why we think that it is meaningful to compare our prediction for a nonpolar homopolymer to experimental results on PNIPAM and PNIPMAM. Our model is not intended to give a detailed description of their phenomenology, though: because of its dimensionality, of its lack of geometric detail, and of the assumptions that make it simple, it is not suited to describe the

“domain” cooperativity within of the coil-globule transition, or the freezing of side chains within the globule.

Warm swelling of PNIPAM and PNIPMAM has not been detected: this is due to the fact that the temperature of warm transition should lie above the water boiling temperature [11].

### III. CONCLUDING REMARKS

We have presented a simple model which reproduces the two-dimensional simulation [5] and the experimental results [8] for the trends of  $\Delta F$ ,  $\Delta E$ ,  $\Delta S$ ,  $\Delta C$  upon hydration of a nonpolar solute, showing that the reduced number of ways of forming hydrogen bonds in the presence of a non-bonding solute, together with a shift in hydrogen bonding energy, are sufficient to reproduce the experimental hallmarks of hydrophobic hydration. This is consistent with the views expressed in Refs. [3,5] and [19].

Our description of water is focused on the clusters in the hydration shell, which are affected by the presence of the solute molecule. To keep the model analytically integrable, we disregard interaction between clusters. In this way, we are able to describe hydration, but cannot describe solvent-mediated long-range interactions between solute molecules. For this reason, this approach is somewhat complementary to the one recently proposed by Kolomeisky and Widom [2], which can be treated analytically in the one-dimensional case, and is very well suited to evaluate the potential of mean force.

When applied to a homopolymer in solution, our model recovers the experimental swelling at both high and low temperatures. We reported only a calculation for a two-dimensional polymer. We did so for the sake of the internal coherence of the model, because the geometric contribution to the density of states of water clusters is estimated for the two-dimensional case and the entropies of both the water and the polymer play a key role in the behavior of the polymer in solution. Anyway, the fact that our two-dimensional water yields a correct qualitative description of nonpolar hydration suggests that the application of the 2-*d* water model to the hydration of 3-*d* SAWs could be interesting; indeed, we have performed some calculations using three-dimensional SAWs and have verified that they also reproduce the same qualitative results; of course, the temperature range of the compact phase varies a bit.

The fact that our model describes both cold and warm swelling of a nonpolar homopolymer suggests that it could be effectively used in protein folding studies, describing both cold and warm unfolding in a single framework. Even if hydrophobicity alone cannot be safely considered as the unique responsible for protein folding and stability, it is very likely that any progress in its description, and in its encoding in simple models, will yield significant improvement in the understanding of the protein

folding puzzle.

### ACKNOWLEDGMENTS

We thank S. Causo, P. De Los Rios and A. Pelizzola for very useful discussions and suggestions. A large part of the present work was carried out when one of the authors (LC) was at the INFM, Politecnico di Torino.

---

\* Electronic address: [pbr@athena.polito.it](mailto:pbr@athena.polito.it).

† Electronic address: [casetti@fi.infn.it](mailto:casetti@fi.infn.it).

- [1] N. Muller, J. Chem. Phys. **43**, 2555 (1965); B. Lee and G. Graziano, J. Am. Chem. Soc. **118**, 5163 (1996).
- [2] A. B. Kolomeisky and B. Widom, Faraday Discuss. **112**, 81 (1999); G. T. Barkema and B. Widom, J. Chem. Phys. **113**, 2349 (2000).
- [3] K. A. T. Silverstein, A. D. J. Haymet, and K. A. Dill, J. Chem. Phys. **111**, 8000 (1999).
- [4] See, e.g., T. E. Creighton, *Proteins, Structures and Molecular Properties* (Freeman, New York, 1993).
- [5] K. A. T. Silverstein, A. D. J. Haymet, and K. A. Dill, J. Am. Chem. Soc. **120**, 3166 (1998).
- [6] J. W. Arthur and A. D. J. Haymet, J. Chem. Phys. **109**, 7991 (1998), and references therein.
- [7] K. A. Dill, Biochemistry **29**, 7133 (1990).
- [8] G. I. Makhataдзе and P. L. Privalov, Adv. Prot. Chem. **47**, 307 (1995).
- [9] C. Wu and X. Wang, Phys. Rev. Lett. **80**, 4092 (1998); E. I. Tiktopulo *et al.*, Macromolecules **27**, 2879 (1994); S. Fujishige, K. Kubota and I. Ando, J. Phys. Chem. **93**, 3311 (1989); A. V. Gorelov, L. N. Vasil'eva, A. Du. Chesne, E. G. Timoshenko, Yu. A. Kunetsov and K. A. Dawson, Il Nuovo Cimento D, **16**, 711 (1994); Y. Tamai, H. Tanaka and K. Nakanishi, Macromolecules **29**, 6750 (1996); Y. Tamai, H. Tanaka and K. Nakanishi, Macromolecules **29**, 6761 (1996).
- [10] X. Wang, X. Qiu and C. Wu, Macromolecules **31**, 2972 (1998).
- [11] G. Graziano, Intl. J. Biol. Macromol. **27**, 89 (2000).
- [12] P. Bruscolini and L. Casetti, Phys. Rev. E **61**, R2208 (2000); P. De Los Rios and G. Caldarelli, *ibid.* **62**, 8449 (2000); **63**, 031802 (2001); A. Bakk *et al.*, cond-mat/0007078; A. Hansen *et al.*, Eur. Phys. J. B **10**, 193 (1999).
- [13] G. M. Bell and D. A. Lavis, J. Phys. A: Gen. Phys. **3**, 568 (1970); A. Ben-Naim, J. Chem. Phys. **54**, 3682 (1971).
- [14] Assuming e.g. that the bond is lost in the “I,E” (rather than in the “E,I”) disposition, there are  $\binom{s}{k}$  ways of choosing the “I”s and  $\binom{m-s}{k}$  of choosing the “E”s for the *k* “I,E” interfaces.
- [15] J. W. Arthur and A. D. J. Haymet, J. Chem. Phys. **110**, 5873 (1999).

- [16] R. Crovetto, R. Fernández Prini and M. L. Japas, J. Chem. Phys. **76**, 1077 (1982).
- [17] J. Douglas C. M. Guttman, A. Mah and T. Ishinabe, Phys. Rev. E **55**, 738 (1997).
- [18] E. E. Gurel N. Kayaman, B. M. Baysal and F. E. Karasz, J. Polym. Sci. B **37**, 2253 (1999).
- [19] N. Matubayasi, J. Am. Chem. Soc. **116**, 1450 (1994).

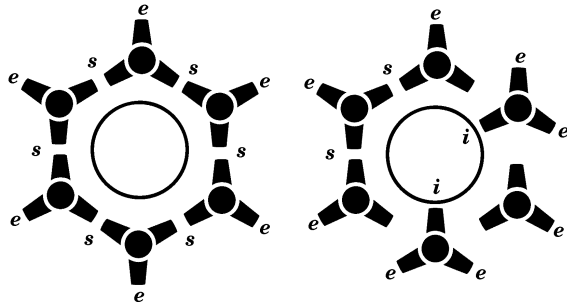


FIG. 1. Cluster of water molecules. Bonds may be formed or broken, according to energy. The circle represents a solute (shell case) or (an)other water molecule(s) (bulk). *Left*: ordered arrangement with molecules in state “E” (one “external” ( $e$ ) and two “on-shell” ( $s$ ) arms). When all bonds are formed this is the ground state. *Right*: in state-I molecules have two  $e$  arms and an internal ( $i$ ) one, which cannot form a bond if a solute is present.

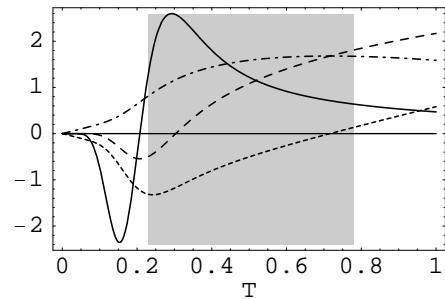


FIG. 2. Free energy  $\Delta F$  (dash-dotted), energy  $\Delta E$  (dashed), entropy  $T\Delta S$  (dotted) and specific heat  $\Delta C/3$  (solid line;  $1/3$  is for rendering purposes) changes upon hydration of a nonpolar solute. Here  $z = 4$ ,  $m = 4$ ,  $q_s = 0$ ,  $q_b = m/2$ ,  $h_s/h_b = 1.2$ ,  $J = 0.3$  (so that  $K = 0$ ). Boltzmann’s constant  $k_B = 1$  so that the specific heat is adimensional. Energies and temperatures are in units of  $h_b$ . The shaded region can be compared with experiments [8] and simulations [5].

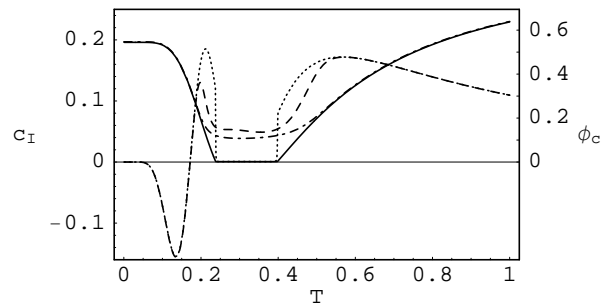


FIG. 3. Nonpolar homopolymer in solution: temperature profiles, for different lengths  $N$ , of the excess specific heat per water site,  $c_I = C_I/N_W$  (left scale; dashed line:  $N = 200$ ; dotted,  $N = 10^6$ ), and average fraction of water-monomer contacts,  $\phi_c = \langle n_c \rangle / (2N + 2)$  (right scale; dash-dotted line:  $N = 200$ ; solid,  $N = 10^6$ ). Parameters as before.